

## Zadatak

---

Napisati parser teksta, *BNFParser*, koji ulazni tekstualni fajl parsira na osnovu modifikovane BNF forme zadate u konfiguracionom fajlu. Izlaz programa treba da bude XML ili JSON fajl (po izboru) koji predstavlja stablo parsiranja ulaznog tekstualnog fajla, na osnovu zadate modifikovane BNF forme.

Modifikovana BNF forma piše se na osnovu pravila za standardnu BNF formu, uz sljedeće dodatke.

- Sa  $\langle a \rangle ::= \text{regex}(\text{regularni izraz})$  označava se čvor koji se određuje na osnovu regularnog izraza specifikovanog u zagradama. U datom slučaju  $\langle a \rangle$  predstavlja bilo koji izraz koji se poklapa sa datim regularnim izrazom.
- Sa  $\langle a \rangle ::= \text{standardni\_izraz}$  označava se čvor koji se određuje na osnovu tabele standardnih regularnih izraza, gdje na mjestu **standardni\\_izraz** može da bude bilo koja od stavki iz sljedeće tabele.

Standardni izraz	Značenje
<b>broj_telefona</b>	Broj telefona, sa ili bez pozivnog broja, sa različitim delimiterima
<b>mejl_adresa</b>	Ispravno formatirana mejl adresa, na proizvoljnom domenu
<b>web_link</b>	Ispravno formatiran apsolutni URL, <a href="https://en.wikipedia.org/wiki/URL">https://en.wikipedia.org/wiki/URL</a>
<b>brojeva_konstanta</b>	Cjelobrojna konstanta ili konstanta sa pokretnim zarezom
<b>veliki_grad</b>	Bilo koji od velikih evropskih gradova (prvih 200), <a href="https://ec.europa.eu/eurostat/web/cities/data/database">https://ec.europa.eu/eurostat/web/cities/data/database</a>

*BNFParser* mora da poštuje sljedeća pravila.

- Konfiguracioni fajl, uvijek nazvan *config.bnf*, mora da sadrži, u modifikovanoj BNF formi, specifikaciju načina parsiranja
- Kao argumente komandne linije, program treba da prihvati ulazni i izlazni fajl, pri čemu izlazni fajl može da bude XML ili JSON formata
- XML tagovi izlaznog fajla (ili *name* polja, u JSON fajlu) moraju biti nazvani na osnovu imena tokena u BNF formi, kako bi rezultujuće stablo parsiranja bilo jasno i ekvivalentno specifikovanoj BNF formi
- U slučaju da ulazni fajl nije moguće parsirati po BNF formi, na standardni izlazni strim ispisati poruku o tome gdje se desila greška (odnosno, liniju iz BNF specifikacije na kojoj se desila greška)

Program je dozvoljeno implementirati u proizvoljnom programskom jeziku, pri čemu nije dozvoljeno korištenje biblioteka za parsiranje i BNF formu. Dozvoljeno je korištenje samo standardne biblioteke odabranog programskog jezika.